PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

#22
6033-04
B.S.H
10f3

In re Application of

ERIC COHEN-SOLAL ET AL

Serial No. 09/488,028

Filed: JANUARY 20, 2000

Atty. Docket US000015

Group Art Unit: 2173

Examiner: T. HAILU

TITLE: MULTI-MODAL VIDEO TARGET ACQUISITION AND RE-DIRECTION SYSTEM AND METHOD

Mail Stop: Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

APPEAL BRIEF

Sir:

REAL PARTY IN INTEREST

The real party in interest is Philips Electronics North America Corporation, having an office at 345 Scarborough Rd., Briarcliff Manor, NY 10510.

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to Mail Stop: Appeal Brief - Patents, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on June 7, 2004.

Daniel E. Tierney, Reg. No. 33,461

RELATED APPEALS AND INTERFERENCES

To the best of Appellant's knowledge and belief, there are no related appeals or interferences.

STATUS OF CLAIMS

Of the original Claims 1-15 filed, Claims 1 and 12-14 were amended, and Claims 3 and 5 were cancelled (without prejudice), by an Amendment mailed September 3, 2002. Also, Claims 1 and 12-14 were amended by an Amendment mailed April 17, 2003. In addition, Claims 1, 8 and 12 were amended by an Amendment sent July 28, 2003. Also, Claims 1 and 12-14 were amended by an Amendment Under 37 C.F.R. 1.116 sent March 5, 2004. Thus, Claims 1, 2, 4 and 6-15 are pending in the Application.

Claims 1, 2, 4, 6-10 and 12-15 stand rejected in the 1/26/04 Final Office Action and Claim 11 is acknowledged as having allowable subject matter. Thus, the rejections of Claims 1, 2, 4, 6-10 and 12-15 are the subject of this appeal.

STATUS OF AMENDMENTS

After the final Office Action of January 26, 2004, the Amendment Under 37 C.F.R. 1.116 was sent on March 5, 2004 including amendments to Claims 1 and 12-14 (as noted above). The Advisory Action of March 18, 2004 indicates that the March 5, 2004 amendments will be entered upon the filing of this Appeal Brief. Thus, the

Appendix to this Appeal Brief includes all pending Claims 1, 2, 4 and 6-15, wherein Claims 1, 8 and 12-14 reflect the amendments referred to in the "Status of Claims" section immediately above, and Claims 2, 4, 6, 7, 9-11 and 15 are as originally filed. [1]

SUMMARY OF THE INVENTION

The claimed invention relates to locating and displaying an image of a target. Among other things, the invention provides various reliable and superior methods of target locating, acquisition, etc. Focusing on Claim 1 as exemplary, Claim 1 recites a method of locating and displaying an image of a target that includes a step of "sensing a triggering event generated by a human operator", where the sensing step includes "sensing a gesture indicating a direction of said target". In one important aspect, Claim 1 recites "receiving additional external information that characterizes at least one machine-sensible feature of a target, said receiving step occurring substantially simultaneously with said sensing step". A camera is aimed "in response to said sensing and said receiving step".

Claim 1 in its entirety recites:

A method of locating and displaying an image of a target, the method comprising the steps of:

sensing a triggering event generated by a human operator;

receiving additional external information that characterizes at least one machine-sensible feature of a target, said receiving step occurring substantially simultaneously with said sensing step; and

---

[1] As noted, Claim 11 was acknowledged by the Examiner as including allowable subject matter and is therefore not at issue in this appeal. It is, however, included in the Appendix for the Board's convenience and because allowance of all pending claims is requested below.

aiming a camera in response to said sensing and said receiving step, wherein said sensing step includes sensing a gesture indicating a direction of said target.

A simple exemplary embodiment that falls within the scope of Claim 1 is instructive of some of its advantages. In the exemplary embodiment (distilled from the description in the specification at pp. 6-8, pp. 14-15 and Fig. 1A), a person points in the direction of a barrel and says the word "barrel". (Specification, p. 6 (lines 28-29), p. 7 (lines 5-8 and 24-26) and Fig. 1A (no. 5)) A first camera of the system recognizes the pointing as a triggering event and may also (but need not) determine a pointing trajectory using the gesture and use it as one input in identifying the target barrel. (Specification, p. 6 (lines 23-27) and p. 15 (lines 21-24)) In addition, the spoken word "barrel" is captured by the system. (Id., p. 7 (lines 5-8)) The spoken word "barrel" is a generic name that describes an object and serves to characterize at least one machine sensible feature of the target barrel: For example, a target "barrel" may be identified via image processing of a video image of an area that includes the barrel. (The images may be captured by a second "target" camera, for example.) Thus, the received word "barrel" may be used to inform the image processing routine what to look for in the image, thus locating the target. (Id., p. 15 (lines 3-5); p. 8 (lines 8-24) The system then aims the second camera at the barrel. (Id., p. 15 (lines 28-30))

It is important to note that as recited in Claim 1 the "additional external information" received "characterizes at least one machine-sensible feature of a target". Thus, in the above example, processing of the speech in itself (for example, via a speech processor) does not necessarily mean that the speech "characterizes at least

one machine-sensible feature of a target". In the example, the speech input received qualifies as "additional external information that characterizes at least one machine-sensible feature of a target" because "barrel" characterizes a target feature that may be captured by a camera and detected by image processing.

Thus, Claim 1 includes a number of advantageous features. For example, receiving additional external information regarding a target substantially simultaneously with the sensing of a triggering event that includes sensing a gesture provides that the additional external information is more reliably identified and more rapidly processed. Also, the additional external information received "characterizes at least one machine sensible feature of a target". Such received information may be utilized to great advantage in locating a target. For example, the above-example demonstrates that the external information received may be correlated with the machine sensible feature of the target, resulting in more flexibility and accuracy in locating the target.

Briefly turning to other aspects of the invention: As indicated in the example above, the gesture included in the triggering event may also be used as one input in identifying the target. Thus, another aspect of the invention includes inputting spatial information to indicate a position of a target where the spatial information includes sensing a gesture indicating a direction of the target. Use of this information together with spoken or other input information about the target provides more accuracy. An exemplary embodiment that includes this aspect is described in the specification at p. 5 (line 16) to p. 6 (line 15). Independent Claim 14 includes recitation relating to this

aspect.

Another aspect of the invention includes scanning for and storing information about potential targets. Such stored information about potential targets is used with a triggering event and additional information about a target in aiming a camera or otherwise locating the target. Exemplary embodiments that include this aspect are described in the specification at p. 11 (line 21) to p. 12 (line 1) and at p. 20 (line 29) to p. 21 (line 17). Claim 12 includes recitation relating to this aspect.

ISSUES

1) Whether each of Claims 1, 2, 4, 6, 7 and 12-15 are anticipated under 35 U.S.C. §102(a) by "The IntelliMedia WorkBench A Generic Environment For Multimodal Systems", by Brondsted et al. ("Brondsted")

2) Whether each of Claims 8-10 are unpatentable under 35 U.S.C. §103(a) over Brondsted in view of "Toward Natural Gesture/Speech HCI: A Case Study Of Weather Narration" by Poddar et al. ("Poddar")

GROUPING OF CLAIMS

1) For the rejection of Claims 1, 2, 4, 6, 7 and 12-15 under 35 U.S.C. §102(a) as anticipated by Brondsted, Claims 1, 2, 4, 6, 7 and 12-15 do not stand or fall together. Rather, (i) Claims 1, 2, 4, 6, 7 and 13, (ii) Claim 12, and (iii) Claims 14 and 15

are each separately patentable groups of claims.

2) For the rejection of Claims 8-10 under 35 U.S.C. §103(a) as unpatentable over Brondsted in view of Podder, Claims 8-10 stand or fall together.

ARGUMENT

For the rejections of Claims 1, 2, 4, 6, 7 and 12-15 as anticipated under 35 U.S.C. §102(a) by Brondsted, (i) Claims 1, 2, 4, 6, 7 and 13 stand or fall together, (ii) Claim 12 stands or falls alone, and (iii) Claims 14 and 15 stand or fall together, because the claims of each group (i), (ii) and (iii) are patentably distinct from the claims of the other groups. Among other things, Claims 1, 2, 4, 6 and 7 (either directly or via dependencies) include recitations relating to at least one "machine sensible" feature as part of a step of receiving additional external information (with analogous recitation for Claim 13), whereas Claim 12 recites various additional steps that include the storing of information concerning machine sensible characteristics and locations of possible targets, which is patentably distinct from Claims 1, 2, 4, 6, 7 and 13. In addition, Claims 14 and 15 (either directly or via dependency) recites input of "spatial information" as well as input of "further information" and orienting an instrument in response to both the spatial and further information "to reduce an ambiguity in said position", which is also patentably distinct from Claims 1, 2, 4, 6, 7 and 13 and from Claim 12.

Claims 8-10 include recitations relating to a "look-up table" and are patentably distinct from the other claims as evidenced by the different basis of rejection cited by

the Examiner.

1) Claims 1, 2, 4, 6, 7 and 13 are not anticipated under 35 U.S.C. §102(a) by Brondsted.

Claim 1 is initially focused on as representative in the ensuing discussion. As noted above, Claim 1 recites among other things "receiving additional external information that characterizes at least one machine-sensible feature of a target". The "additional external information" received in Claim 1 is "additional" to sensing a triggering event that includes sensing a gesture indicating a direction of the target. As exemplified above, the "additional external information" recitation of Claim 1 may include speech input, for example. In that case, Claim 1 requires that the speech input "characterizes at least one machine-sensible feature of a target".[2]

In rejecting Claim 1, the Examiner points to Sections 2.1 and 3 of Brondsted as purportedly teaching the "receiving additional external information" recitation of Claim 1. (1/26/04 Final Office Action, ¶3 at pp. 2-3) The Examiner also refers to Sections 2.1 and 2.2 of Brondsted in responding to Appellants arguments (which are presented below).[3] (1/26/04 Final Office Action, ¶6)

Although Brondsted refers to both gesture and speech recognition, the claimed

---

[2] As also emphasized above in the Summary section, this recitation of Claim 1 is not directed at whether the external information itself is machine sensible, but rather recites that the external information "characterizes" at least one machine sensible feature of a target.

[3] In the Advisory Action of March 18, 2004, the Examiner repeats the citation to Section 3 of Brondsted, and includes a non-specific reference to "elsewhere" in Brondsted. It is not clear whether the Examiner is again referring to Sections 2.1 and 2.2. Be that as it may, Appellants thoroughly present and distinguish herein the pertinent aspects of Brondsted.

invention is clearly distinguishable. The application presented in Section 2.1 of Brondsted is a "Campus Information System" where a user, for example, inquires about routes from one location to another, where a given person's office is located, etc. Brondsted describes a user querying the system by speech input (such as "Show me Hanne's office") as well as gesture input (such as "pointing coordinates"). However, the spoken query inputs in Section 2.1 of Brondsted (such as "Show me Hanne's office") are not described by Brondsted as characterizing a machine-sensible feature of a target. Thus, the spoken query inputs in this section of Brondsted do not teach "receiving additional external information that characterizes at least one machine-sensible feature of a target", as recited in Claim 1.

As to Section 3 of Brondsted, the gesture and speech inputs related to the Campus Information System are consistent with those of Section 2.1 discussed above. The speech inputs described in Section 3 are "routes", "room numbers" and "names" of a person.[4] (Brondsted, Section 3, "Speech Recognizer" and "Domain Model") Section 3 makes it even clearer than Section 2.1 that the speech input does not teach "receiving additional external information that characterizes at least one machine-sensible feature of a target": The description of the Domain Model is an information

---

[4] The only specific references to names in Bronsted are names of persons. Domain Model of Section 3 also receives "coordinate" input, which Brondsted clearly ties to gesture input, not speech. (Brondsted, Section 2.1 ("Frame semantics", referring to gesture inputs as "e.g. pointing coordinates") and Section 3.0 ("Gesture Recognizer", describing determining the position the user is pointing to))

database used for straight-forward processing that simply associates the spoken inputs with locations (among other outputs).[5] Brondsted does not teach any sensors or the like that detect features of a target which are characterized by these spoken inputs. In addition, there are no features in the Campus Information System as shown in Fig. 1 that could even be detected by a sensor or the like that is characterized by a spoken "name" of a person, "route" or "room number". Thus, the spoken inputs of the Campus Information System of Section 3 of Brondsted also do not teach "receiving additional external information that characterizes at least one machine-sensible feature of a target".[6]

The application presented in Section 2.2 of Brondsted ("Automatic Pool Trainer") is also far afield from the claimed invention. In this application, when the user points the cue toward the cue ball, the system calculates and displays the trajectories of the pool balls based on the direction of the cue. There is also reference to the user issuing spoken commands to the system, without any further elaboration apart from the Section 3 statement that they are "keywords". Thus, even assuming arguendo (but not conceding) that the cue ball is considered a "target" and the cue provides a gesture

[5] Section 3 indicates that the Domain Model processes the "coordinate" input in the same manner.

[6] Although not taught by Brondsted, there are circumstances where, for example, a "room number" (or a person's "name", etc.) might characterize at least one machine-sensible feature of a target and thus be encompassed by Claim 1. For example, an area having a number of doors displaying their room numbers that could be recognized via optical detection. In such context, speaking a "room number" or like input would fall within the scope of "receiving additional external information that characterizes at least one machine-sensible feature of a target".

indicating a direction of the target, there is no teaching that the command or keyword "characterizes at least one machine-sensible feature of a target" as recited in Claim 1.

Thus, the Campus Information System, the Automatic Pool Trainer, and the Domain Model related thereto of Brondsted do not teach "receiving additional external information that characterizes at least one machine-sensible feature of a target".

In order to make a proper anticipation rejection under 35 U.S.C. §102(a), MPEP 2131 requires that a reference must show the identical invention. Because Brondsted fails to show at least the above-noted recitations of independent Claim 1, Brondsted fails to anticipate Claim 1.

Independent Claim 13 recites "inputting further information about a machine-sensible characteristic of said target". Independent Claim 13 thus contains analogous distinguishing recitation as discussed above for Claim 1 and is distinguished from Brondsted for like reasons.

For at least the above-noted reasons, Brondsted fails to anticipate independent Claims 1 and 13. Dependent Claims 2, 4, 6 and 7 are distinguished from Brondsted by virtue of their dependencies on independent Claim 1. Allowance of Claims 1, 2, 4, 6, 7 and 13 is respectfully requested.

2) Claim 12 is not anticipated under 35 U.S.C. §102(a) by Brondsted.

Turning to independent Claim 12, the "receiving additional external information" step recited in Claim 12 does not include recitation relating to a "machine sensible" feature

of the target as in Claim 1 (and like recitation in Claim 13). However, Claim 12 includes steps of:

> "scanning an area within the range of at least one sensor;
> identifying potential targets;
> storing information concerning machine sensible characteristics and locations of said possible targets".

In addition, as recited in Claim 12, a camera is aimed in response to the sensing of a triggering event, the storing of information concerning machine sensible characteristics and locations of possible targets, and receipt of additional external information characterizing at least one feature of a target. Among other things, such storage of information as recited in Claim 12 allows a target to be more rapidly located once the triggering event and additional external information is received.

Brondsted fails to anticipate Claim 12: The Campus Information System describes speech input (as well as gesture based input) that is used in a "Domain Model" to generate a corresponding output. (Brondsted, Section 3, "Domain Model") Thus, this aspect of Brondsted fails to teach at least the scanning, identifying and storing steps recited in Claim 12, as well as the aiming of a camera in response to the storing, sensing and receiving steps as recited in Claim 12. For the Automatic Pool Trainer application of Brondsted, the system locates the positions of the pool balls on the table before the cue is pointed. (Brondsted, Section 2.2) However, as noted in part 1 above, the user issues spoken commands to the system, without any further elaboration apart from the Section 3 statement that they are "keywords". Thus, this aspect of Brondsted fails to teach at least the Claim 12 recitation of "receiving

additional external information that characterizes at least one feature of said target".

For at least the above-noted reasons, Brondsted fails to teach the identical invention as recited in Claim 12 and thus fails to anticipate independent Claim 12.[7] Allowance of Claim 12 is respectfully requested.

3) Claims 14 and 15 are not anticipated under 35 U.S.C. §102(a) by Brondsted.

Independent Claim 14 does not include steps analogous to the scanning, identifying and storing steps included in Claim 12; nor does Claim 14 otherwise include a "machine sensible" recitation. However, independent Claim 14 specifically recites "inputting spatial information to indicate a position of a target", where the spatial information "includes sensing a gesture indicating a direction of said target". Claim 14 also recites "inputting further information about said target" which may, for example, comprise speech input. Claim 14 acquires a target using both inputs "to reduce an ambiguity in said position" of the target. As noted, the two inputs improve reliability.

Brondsted fails to anticipate Claim 14. As detailed in part 1 of the argument above, Brondsted's Campus Information System shows speech input that is only used in a "domain model" to generate a corresponding output. The spoken inputs for the

[7] For Claim 12, Brondsted also does not suggest that the missing aspects of the Automatic Pool Trainer be gleaned from the Campus Information System, or vice versa: Brondsted specifically states that the Automatic Pool Trainer does not use the Natural Language Parser, due to the simple mapping from keywords to commands. (Brondsted, Sec. 3, "Natural Language Parser") Also, locating the positions of the pool balls as in the Automatic Pool Trainer is clearly inapplicable to the Campus Information System, which relies on a Domain Model having fixed information stored in a database. (Brondsted, Section 3, "Domain Model")

Automatic Pool Trainer are "commands". Although the input "further information" of Claim 14 may comprise speech, there is nothing in Brondsted that teaches orienting an instrument in response to input spatial information (that includes sensing a gesture indicating a direction of a target) and input further information *to reduce an ambiguity in the position of the target*. For at least these reasons, Claim 14 is not anticipated by Brondsted. Dependent Claim 15 is distinguished from Brondsted by virtue of its dependency on Claim 14. Allowance of Claims 14 and 15 is respectfully requested.

4) Claims 8-10 are not unpatentable under 35 U.S.C. §103(a) over Brondsted in view of Poddar.

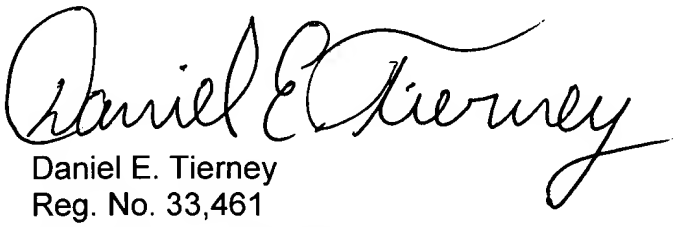
Dependent Claims 8-10 were rejected in paragraph 4 of the 1/26/04 Final Office Action as unpatentable over Brondsted in view of Poddar. Poddar relates to display control and human computer interface (e.g., Poddar section 1). Poddar principally presents a study of particular keywords that appear to co-occur with certain gestures in weather narration, and suggests that keywords be used to more accurately identify gestures in themselves (Poddar, Section 3). The Examiner refers to the experimental results of this study given in Tables 1-4 of Poddar as purportedly showing the "look-up tables" as recited in Claims 8-10. The Examiner also refers to Section 3 of Poddar as disclosing a multi-modal system comprised of speech and gesture input. (1/26/04 Final Office Action, ¶4)

As presented in part 1 of the argument above with respect to independent Claim 1, upon which Claims 8-10 depend, Brondsted fails to teach at least the Claim 1

recitation of "receiving additional external information that characterizes at least one machine-sensible feature of a target".[8] The Examiner presents nothing in paragraph 4 of the Office Action from Poddar to cure this deficiency of Brondsted, and it is submitted that the teachings of Poddar either do not cure the deficiencies of Brondsted with respect to Claim 1, or that the context of the teachings fail to provide a proper motivation or suggestion for combining them with Brondsted. Thus, Brondsted in view of Poddar fails to present a prima face case of obviousness at least under MPEP 2143.01 or 2143.03 with respect to Claim 1, upon which Claims 8-10 depend. Thus, Claims 8-10 are allowable and allowance is respectfully requested.

SUMMARY

For the reasons given above, each of Claims 1, 2, 4, 6-10 and 12-15 are allowable. As noted in footnote 1, dependent Claim 11 is acknowledged by the Examiner as having allowable subject matter. Thus, allowance of Claims 1, 2, 4 and 6-15 is respectfully requested.

By: 

Daniel E. Tierney
Reg. No. 33,461
Attorney For Applicant
631-588-4429

Dated June 7, 2004

[8] For the purposes of Claims 8-10, it is clear from the discussion of Brondsted that Brondsted also fails to suggest all of the recitations of Claim 1.

APPENDIX

CLAIMS ON APPEAL

1. A method of locating and displaying an image of a target, the method comprising the steps of:

sensing a triggering event generated by a human operator;

receiving additional external information that characterizes at least one machine-sensible feature of a target, said receiving step occurring substantially simultaneously with said sensing step; and

aiming a camera in response to said sensing and said receiving step, wherein said sensing step includes sensing a gesture indicating a direction of said target.

2. The method of claim 1, wherein said sensing step includes sensing a gesture of a human operator indicating a target.

3. (Canceled)

4. The method of claim 1, wherein said receiving step includes receiving speech from said human operator.

5. (Canceled)

6. The method of claim 4, further including processing said speech for use with at least one machine sensor, said at least one machine sensor and said speech assisting in locating said target.

7. The method of claim 6, wherein said sensing step includes sensing a gesture indicting a direction from said human operator to said target.

8. The method of claim 6, wherein said processing step includes processing said voice information through a look-up table corresponding said speech to search criteria for use with said at least one sensor.

9. The method of claim 8, wherein said look-up table is modifiable.

10. The method of claim 9, wherein said look-up table is modified by receiving information through the on-line global computer network.

11. (Original -- Allowable) The method of claim 9, wherein said look-up table is modified to include an additional voice input and a corresponding search criteria, said added voice input and said corresponding search criteria established by comparing previous association of said added voice input with at least one machine sensible characteristic of at least one correctly identified target associated with said voice input,

said machine sensible characteristic being a basis for determining said corresponding search criteria.

12. A method of locating and displaying an image of a target, the method comprising the steps of:

scanning an area within the range of at least one sensor;

identifying potential targets;

storing information concerning machine sensible characteristics and locations of said possible targets;

sensing a triggering event, said triggering event generated by a human operator;

receiving additional external information that characterizes at least one feature of said target, said receiving step occurring substantially simultaneously with said sensing step; and

aiming a camera in response to said sensing, storing and said receiving steps, wherein said sensing step includes sensing a gesture indicating a direction of said target.

13. A method of aiming a camera at a target, comprising the steps of:

inputting an indication of a position of a target;

inputting further information about a machine-sensible characteristic of said target;

aiming a camera at said target in response to said indication and said further information to reduce an error in said aiming, wherein said inputting an indication step includes inputting a gesture indicating a direction of said target.

14. A method of acquiring a target, comprising the steps of:

inputting spatial information to indicate a position of a target;

inputting further information about said target; and

orienting an instrument with respect to said target to acquire said target in response to said spatial information and said further information to reduce an ambiguity in said position, wherein said spatial information includes sensing a gesture indicating a direction of said target.

15. A method as in claim 14, wherein said step of orienting includes orienting a camera.

120 AF 2173 /B

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re Application of	Atty. Docket
ERIC COHEN-SOLAL ET AL	US000015
Serial No. 09/488,028	Group Art Unit: 2173
Filed: JANUARY 20, 2000	Examiner: T. HAILU

TITLE: MULTI-MODAL VIDEO TARGET ACQUISITION AND RE-DIRECTION SYSTEM AND METHOD

SUBMISSION OF APPELLANT'S BRIEF ON APPEAL

Mail Stop: Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Submitted herewith please find an original and two copies of Appellant's Brief on Appeal. A check in the amount of $330 is enclosed for the statutory fee under Rule 1.17(c).

Respectfully submitted,

By Daniel E. Tierney
Daniel E. Tierney
Reg. No. 33,461
(631) 588-4429

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to Mail Stop: Appeal Brief - Patents, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on June 7, 2004.

Daniel E. Tierney
Daniel E. Tierney, Reg. No. 33,461